Journal of Environmental Management 75 (2005) 285–301

Journal of Environmental Management

www.elsevier.com/locate/jenvman

Estimation of biogenic emissions with satellite-derived land use and land cover data for air quality modeling of Houston-Galveston ozone nonattainment area

Daewon W. Byun[a,*], Soontae Kim[a], Beata Czader[a], David Nowak[b], Stephen Stetson[c], Mark Estes[d]

[a]*Institute for Multi-dimensional Air Quality Studies, University of Houston, Houston, TX 77204-5007, USA*
[b]*USDA Forest Service, Northeastern Research Station, Syracuse, NY, USA*
[c]*Global Environmental Management, Camden, ME, USA*
[d]*Texas Commission on the Environmental Quality, Austin, TX, USA*

Received 10 May 2004; revised 16 October 2004; accepted 23 October 2004

Abstract

The Houston-Galveston Area (HGA) is one of the most severe ozone non-attainment regions in the US. To study the effectiveness of controlling anthropogenic emissions to mitigate regional ozone nonattainment problems, it is necessary to utilize adequate datasets describing the environmental conditions that influence the photochemical reactivity of the ambient atmosphere. Compared to the anthropogenic emissions from point and mobile sources, there are large uncertainties in the locations and amounts of biogenic emissions. For regional air quality modeling applications, biogenic emissions are not directly measured but are usually estimated with meteorological data such as photo-synthetically active solar radiation, surface temperature, land type, and vegetation database. In this paper, we characterize these meteorological input parameters and two different land use land cover datasets available for HGA: the conventional biogenic vegetation/land use data and satellite-derived high-resolution land cover data. We describe the procedures used for the estimation of biogenic emissions with the satellite derived land cover data and leaf mass density information. Air quality model simulations were performed using both the original and the new biogenic emissions estimates. The results showed that there were considerable uncertainties in biogenic emissions inputs. Subsequently, ozone predictions were affected up to 10 ppb, but the magnitudes and locations of peak ozone varied each day depending on the upwind or downwind positions of the biogenic emission sources relative to the anthropogenic NO_x and VOC sources. Although the assessment had limitations such as heterogeneity in the spatial resolutions, the study highlighted the significance of biogenic emissions uncertainty on air quality predictions. However, the study did not allow extrapolation of the directional changes in air quality corresponding to the changes in LULC because the two datasets were based on vastly different LULC category definitions and uncertainties in the vegetation distributions.

Keywords: Biogenic emissions; Land use/land cover; Satellite data; Air quality; Ozone; Leaf mass density; Photosynthetically active solar radiation; TexAQS experiment

1. Introduction

Development of reliable land use (LU) and land cover (LC) data is a key activity of environmental management. The LULC data is a critical input for characterizing land surface processes in meteorological modeling and in estimating the influence of trees on urban and regional air quality. Different distributions of LULC can influence meteorological conditions such as air temperature, moisture, and the height of the planetary boundary layer in which pollutants are dispersed and transformed. Changes in land

* Corresponding author. Address: Department of Geosciences, Institute for Multi-dimensional Air Quality Studies, University of Houston, 312 Science and Research Building 1, Houston, TX 77204-5007, USA. Tel.: +1 713 743 0707; fax: +1 713 748 7906.

E-mail addresses: dbyun@mail.uh.edu (D.W. Byun), dwbyun@math.uh.edu (D.W. Byun).

0301-4797/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jenvman.2004.10.009

use and land cover in a rapidly developing metropolitan area can modify radiative conditions and thus result in regional climate changes. For example, LULC modification can exacerbate the urban heat island (UHI) effects in a metropolitan area like Houston-Galveston when the vegetation cover is replaced with buildings and pavement. Also, the extended impervious surface may enhance the run off of rainfall and increase the sensible heat flux and air temperature, thus modifying the turbulent mixing conditions in urban areas (Oke, 1987; Taha et al., 2000). There are a few attempts to mitigate such problems by introducing more plants in the urban environment, paving with light color and porous materials, and installing reflective roofing materials.

Although reducing the effects of the UHI can be beneficial, for example by reducing summer air conditioning costs, there is a controversy whether these measures will always mitigate local ozone problems. Changes in the land use and land cover will induce further changes in the amount vegetation can eliminate pollutants from the air through dry deposition. Deciduous trees are well known to generate a significant amount of photochemically sensitive biogenic emissions of volatile organic compounds (VOCs), carbon monoxide (CO) and nitrogen monoxide (NO), as well as monoterpenes which are precursors of secondary organic aerosols (SOAs). The rates of the biogenic emissions are heavily dependent on the environmental factors that affect tree and vegetation physiology, such as ambient temperature, moisture, and in particular the photosynthetically active solar radiation (PASR) component. For example, lower temperature could have beneficial effects by reducing the amount of such ozone precursor emissions. Refer to Guenther et al. (1993, 1994) and Pierce et al., (1998) for the details. However, decreased temperatures could also produce lower mixing heights in which precursors and secondary air pollutants can disperse, thus increasing ozone concentrations. Therefore, there is a need to perform a thorough assessment of these issues by using advanced computer modeling to evaluate whether an expanded tree planting program would mitigate ozone problems in an urban area like the Houston-Galveston Area (HGA).

An LULC dataset is an essential input for the estimation of regional biogenic emissions. The higher the spatial resolution and the larger the numbers of plant species available, the better the estimations are that can be expected. Texas Commission on Environmental Quality (TCEQ, 2002) utilizes a high quality LULC dataset (Wiedenmyer et al., 2001) for biogenic emission estimates. However, there are a few areas of uncertainties with the LULC dataset because it is a compilation of various data sources over many years of time span and has many components that depend on the county-based surrogate information (for example population, socio-economic databases, crop yields for each count, etc.), and therefore, some spatial distributions are limited by county boundaries. Recently, the Texas Forest Service (TFS), with the support of TCEQ, has compiled a new high-resolution LULC dataset for the eight counties surrounding the HGA to characterize regional changes in the vegetation and tree species. The updated map of LULC was produced using satellite imagery and ancillary datasets. A supervised classification process using image processing software was employed to define 8 land cover classes and 15 land use classes (GEM, 2003).

The new set of vegetation and LULC data, when compared with the TCEQ dataset, provides an opportunity to evaluate the impact of LULC data uncertainties on the biogenic emission estimates and eventually on air quality simulations. The present research examines the effects of LULC changes on the local biogenic emission estimates without considering the indirect effects on the meteorological conditions. In this study two different biogenic emission datasets are used in air quality simulations to quantify their effects on the predicted ozone concentrations in the HGA.

2. Materials for analysis

The study area for land use and land cover is the Houston-Galveston area including the surrounding counties of Brazoria, Chambers, Fort Bend, Galveston, Harris, Liberty, Montgomery, and Waller in South East Texas, as shown in Fig. 1. The period of modeling analysis is August 22–September 1, 2000, a part of the Texas Air Quality Study 2000 (TexAQS 2000).

2.1. Land cover and land use data

2.1.1. TCEQ biogenic landuse data set

For the HGA State Implementation Plan (SIP), TCEQ supported development of a dataset of land use and vegetation for the state of Texas and the surrounding states. Compared to other parts of Texas and the US, the LULC database available for Eastern Texas has been updated relatively recently (Wiedenmyer et al., 2001). It is a composite land use database that includes a mapping of ground cover, vegetation species, and leaf mass densities for the state of Texas. Land use and vegetation were divided into over 600 classifications at approximately 1 km spatial resolution (Fig. 2). Some field surveys were performed to estimate leaf biomass densities of certain tree species. The dataset contains more detailed urban LULC classifications. When no recent data were available, the USGS-LULC database at 1 km resolution (http://www.lib.ncsu.edu/stacks/gis/lulc.html) was applied to provide spatial distribution of the urban land use types. In addition to the municipal, state, and Federal government land use, land cover, and vegetation data at resolutions from 30 m to 1 km, county-based agricultural LULC data were incorporated as well. Although this LULC database has a reference year of

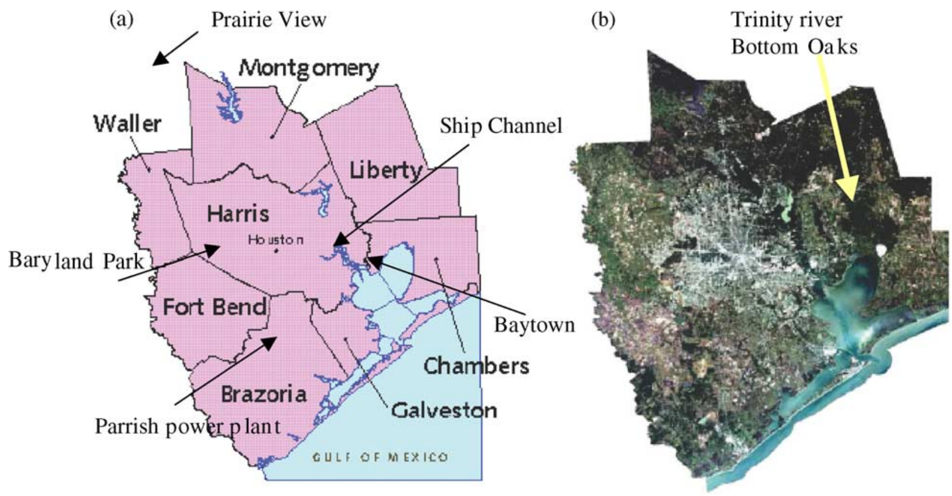

Fig. 1. (a) The study area consisting of the eight counties around Houston, Texas, including Brazoria, Chambers, Fort Bend, Galveston, Harris, Liberty, Montgomery, and Waller. (b) Corresponding LANSDAT image (from GEM, 2003).

2000, the representative years of the different data sources vary widely and in some cases are uncertain. Hereafter, we will call this dataset TCEQ-LULC.

To estimate biogenic emissions, species distributions and biomass densities were associated with each land use type based on the field surveys that took place throughout Central and Eastern Texas between 1997 and 1999 (Wiedenmyer et al., 2001; Funk et al., 2001). Urban land areas near Dallas-Fort Worth, Austin, Beaumont-Port Arthur, and Houston-Galveston were surveyed and tree size and species distribution information was collected. Biomass density was estimated for each species for each land use type using the empirical relations suggested by Geron et al. (1994). Each LULC type was assigned a five-digit code and the biomass densities and species distribution were assigned to each code. According to Wiedenmyer et al. (2001), the total leaf biomass densities associated with these codes for Texas range from 0 to 556 g m^{-2} with the highest value being in central and eastern Texas where the high isoprene-emitting oak species is most abundant. This dataset can be considered as more detailed and having updated land type distributions than the BELD3 dataset available from EPA (Kinnee et al., 1997). BELD3 is used to estimate biogenic emissions across the conterminous US continent. The TCEQ dataset is available on the web from http://www.tnrcc.state.tx.us/air/aqp/airquality_photomod.html#ei4c.

2.1.2. LANDSAT based landuse data set

The TCEQ biogenic landuse dataset used for the HGA SIP is one of most advanced ones. However, there are a few areas of uncertainties with the data such that (1) the TCEQ-LULC is a compilation of various data sources over many years of time span, (2) it has many components that depends on the county-based surrogate information, such as population and other socio-economic databases (such as crop yields for each county), and therefore, some spatial distributions are limited by county boundary.

Recently, Texas Forest Service (TFS) has prepared a detailed land cover and land use data set from LANDSAT-derived data that is based on multi-spectral land-surface characteristics for the purpose of managing the urban forest in the HGA. Therefore, it can serve as an independent dataset to verify the TCEQ-LULC; then we can use it to quantify uncertainties in the TCEQ-LULC data. Because biogenic emissions are one of very important inputs determining the background (natural) conditions for ozone air quality, it is beneficial to study the uncertainty in the biogenic emissions by comparing the two different land use datasets.

The land use land cover GIS shape files (GEM, 2003) cover the eight county areas surrounding Houston (see Fig. 1) and consists of 8 land cover classes and 15 land use classes. An updated map of Land Use and Cover Type conditions within the eight counties surrounding Houston, Texas, was produced using satellite imagery and ancillary data sets (Fig. 3). A supervised classification process using image processing software was employed to define the classes described below:

Land use: forest, range, agriculture, urban/developed
Land cover: forest composition (coniferous, broadleaf, mixed), grass, wet land, water, barren (e.g. beach, bare soil), impervious (roads, parking lots, buildings)

Urban and non-urban areas were separated with year 2000 census data, and land use. Land use conditions within the urban areas were further classified as residential, commercial, industrial, transportation corridors, and parks. Surface properties including albedo, surface roughness, emissivity, and thermal inertia were mapped through direct measurements when possible and otherwise through look-up-tables that assign a value according to cover type. Several data sources were utilized to enhance feature identification and to improve the classification process.

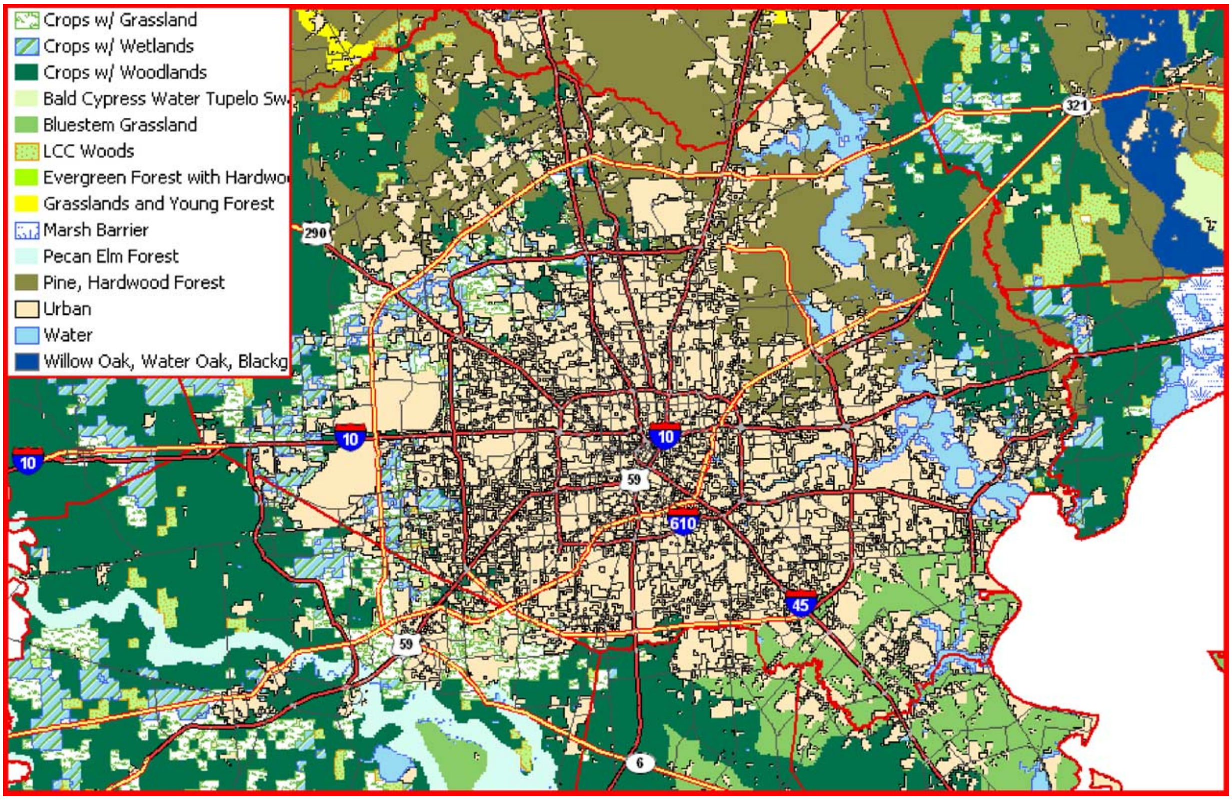

Fig. 2. TCEQ biogenic land use dataset used for the modeling of the TexAQS 2000 episode. The original land cover classification of 600 types is aggregated to 14 vegetation and land cover types for readability.

The data sources included LANDSAT satellite imagery, digital aerial photography, aerial LIDAR data, USGS elevation data, a variety of vector based GIS data, and extensive ground-truth information obtained from more than 300 field plots.

The multi-resolution and multi-temporal datasets were correlated to develop unique class parameters for optimal recognition of land use and cover type features. The higher resolution datasets were used to perform a sub-pixel analysis of the LANDSAT imagery over selected field plot locations in order to produce spectral signatures for the associated cover types. The spectral signatures were then extrapolated across the entire study area to assign a land cover type to each picture element (pixel) in the satellite imagery using the Spectral Angle Mapper algorithm within the ENVI software package by Research Systems, Inc.

The land use classification was conducted following the cover type analysis by means of a manual interpretation of all the data, including the Cover Type classification, the LANDSAT imagery and the ancillary datasets. On screen digitizing of land use class boundaries was performed for the urban sub-classes of residential, commercial, industrial and parks. The total volume of data for the project exceeded a terabyte of information. Datasets in raster format included LANDSAT 7 satellite data, aerial photography, LIDAR data and USGS elevation data. Datasets in vector format included field plot data, county boundaries, Census 2000 urban boundaries, transportation routes, hydrology, wetlands, parks (City, County, State and National), and grid domains for air quality modeling activities. Hereafter, we will call this LANDSAT based data as the TFS-LULC.

It should be noted that the new TFS-LULC data may represent neither the identical reference year nor the changes in the forest cover in Houston because the original TCEQ-LULC data is a composite of several different datasets completed during different years. The differences we see between the two are probably due to the changes in forest density, though there may be other factors.

2.2. Meteorological data

Meteorological data needed for air quality assessment studies are usually prepared using a sophisticated numerical weather model, often with some kind of measurement data assimilation techniques to best characterize past meteorological conditions. Data from meteorological models are used in photochemical models to characterize advection, dispersion, cloud mixing, and dry and wet deposition processes. Recent studies suggest that the peculiar meteorological conditions in the HGA, such as land and sea breeze, development of the coastal internal boundary layer, urban heat island phenomena, and the convergent influx of humid and warm Gulf air mass into the region all affect the formation, accumulation, and transport of ozone and other

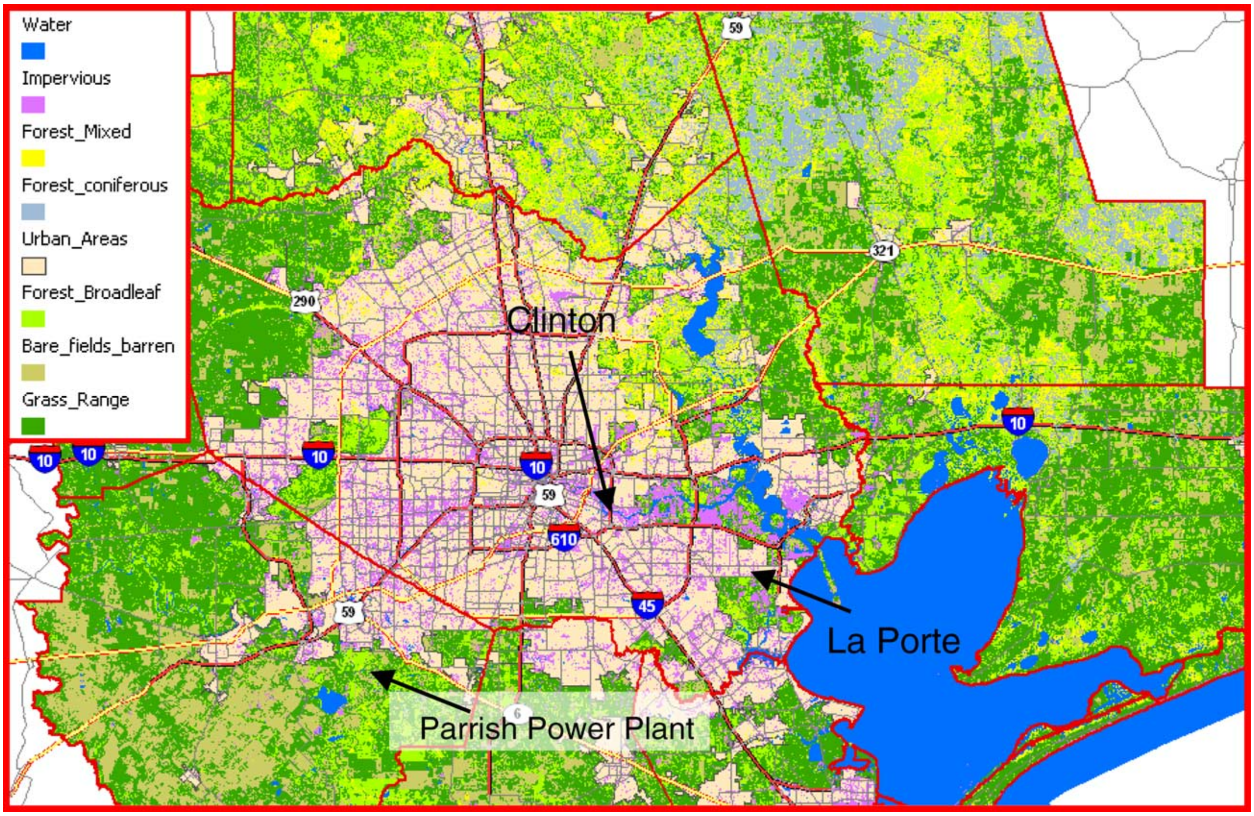

Fig. 3. A composite map of land cover and urban land use classification derived from the LANDSAT data. The urban category here includes urban commercial and residential areas.

pollutants in very complex ways (Nielsen-Gammon, 2001; Nielsen-Gammon et al., 2004). Further, our preliminary modeling study indicates that meteorological simulation results for the HGA are very sensitive to the land surface characterization data (e.g. land use and soil moisture), model vertical grid structure, and algorithms used for the parameterization of the planetary boundary layer (e.g. Byun et al., 2003).

Use of different land use inputs produces distinctively different meteorological simulation results. In addition to the direct impact of the leaf mass density and vegetation type information pertaining to the LULC inputs, changes in the temperature, cloudiness, and radiation conditions also affect the biogenic emissions rate estimates. Therefore, to fully account for the effects of the LULC on the air quality, the new TFS-LULC data must be used to characterize the land surface processes in the meteorological model, and the new meteorological model output must be used to provide inputs for the biogenic and anthropogenic emission estimates. Then air quality model simulations need to be done using all these new inputs. While the research is continuing to assess these cascading impacts of land use change on the air quality (e.g. Nowak et al., 2000), in this study, we have utilized the observation based temperature and PASR inputs for estimating biogenic emissions. This practice allows estimation of the biogenic emissions more realistically and also helps in studying the direct effects of using the new land use data for the biogenic emissions estimates in isolation. However, air quality model simulations are performed with the MM5 simulations results, which are described below.

2.2.1. Observation based temperature data

Temperature measurements were obtained from several different monitoring networks. Networks were chosen if they had acceptable quality assurance procedures in place and if data were available for the time period of interest. Differences in sensor height among the temperature networks are usually not an issue during hot summer days, when vigorous mixing leads to small temperature gradients, but they might be an issue during dry, cool, still conditions when larger temperature gradients might occur near the ground. Data from the following networks were used: Texas Commission on Environmental Quality network, Aerometric Information Retrieval System network, National Weather Service network, Texas Crop Weather Program, Conrad Blucher Institute Texas Coastal Observation Network, and National Automated Buoy Data network. Overall, data from over 100 stations were used.

The statistical technique of kriging was used to interpolate temperature measurements, thus creating a temperature field for each hour of the chosen episode. Vizuete et al. (2002) found that kriging is one of the most effective temperature interpolation methods for creating biogenic emission model inputs. Kriging takes into account the tendency of neighboring observations to be more alike than those that are far apart. The function that describes the average similarity of any two observations as a function of distance is called the semivariogram. Because there was considerable variability in the semivariograms calculated for different times during the day, unique semivariograms were estimated for each hour. Specifically, a power function was fitted to each hourly semivariogram, and the fitted power function was used in the kriging algorithm. Therefore, each hour had a different semivariogram as the basis of the interpolation. The SAS© software kriging algorithm was used in this application. Temperature fields were calculated for each hour at three different spatial resolutions: 4 km×4 km, 12 km×12 km, and 36 km×36 km grid cells. The different grids were nested within each other, and were configured to exactly match the photochemical modeling domains.

Data from a temperature site not used in the interpolations were compared to the temperature field values at that location. The Soil Climate Analysis Network (SCAN) data site was located at Prairie View A&M University, in Waller County, Texas. Fig. 4 shows a sample time series of the interpolated temperatures and the measured temperatures. The time series indicates that in general, the interpolated temperatures usually depict the diurnal variation of temperature at the site reasonably well. It also shows that the overnight temperatures were generally overestimated, and the maximum temperatures, especially on very hot days, were sometimes underestimated. A scatter plot of the same data (Fig. 5) shows a high degree of correlation ($r^2=0.94$) between the measured and modeled values. The 1:1 line indicates that the interpolation overestimates temperatures on the low end, but generally depicts the higher temperatures (i.e. >30 °C) well. Since the higher temperatures are more important in biogenic emissions estimates, the temperature interpolation seems to be a sound method for estimating temperatures for biogenic emissions modeling.

2.2.2. Photosynthetically-active solar radiation data

Photosynthetically-active solar radiation (PASR) is defined as visible radiation with wavelengths between 400 and 700 nm. Biogenic emissions modeling requires the input of hourly PASR fields that extend over large domains. Interpolation of surface measurements is unlikely to yield a satisfactory field, given the heterogeneous nature of clouds, and the comparative rarity of PASR measurements. Meteorological models can generate PASR fields, but sometimes generate spurious clouds, which greatly affect the PASR field. Therefore, hourly PASR fields were created using algorithms developed by Pinker et al. (2003) with input data from the GOES8 satellite. Cloud cover estimates from satellite imagery were fed into the radiation balance algorithm to create a large-scale field of PASR. High resolution PASR fields were created from 1/16° solar field data.

Comparisons between GOES-derived PASR fields and ground-based broadband solar radiation measurements found very high degrees of correlation. Correlations for TCEQ sites ranged from 0.94 to 0.97, with slopes ranging from 0.47 to 0.53, indicating that PASR comprised approximately 50% of broadband solar radiation (i.e. 20–2000 nm). Fig. 6 shows an example of the time series comparison between the GOES-derived PSAR values and the broadband solar measurements at the Baryland Park TCEQ solar radiation site (see Fig. 1). The nearest measurements of PASR at a ground station were in

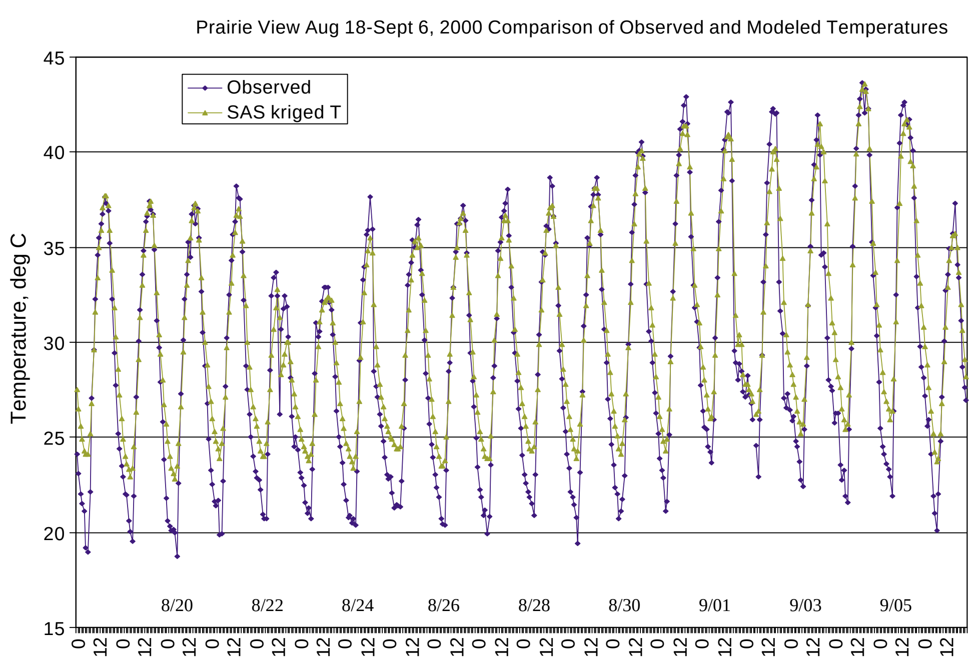

Fig. 4. A sample time series of the interpolated temperatures and the measured temperatures at the Prairie View for August 19–September 6, 2000.

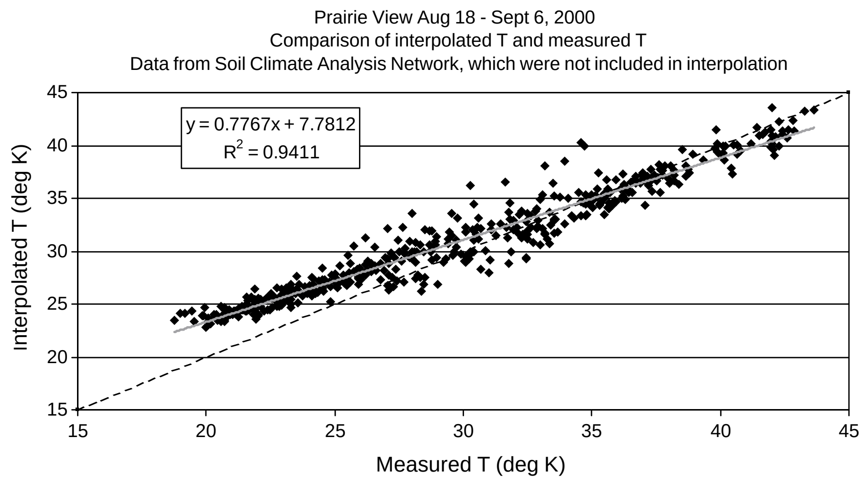

Fig. 5. A sample scatter plot of the interpolated temperatures versus the measured temperatures at the Prairie View for August 19–September 6, 2000.

Mississippi at the CONFRRM monitoring sites located in that state. Since those sites are outside the 4 km domain, the comparisons of GOES-derived data and ground observations are not as useful.

2.2.3. MM5 simulation data

In support of the TCEQ's SIP modeling studies, Dr John Nielsen-Gammon of Texas A&M University (TAMU) simulated the NCAR/Penn State (National Center for Atmospheric Research/the Pennsylvania State University) Mesoscale Model, Version 5, Release 3.6 (MM5V3.6) (Grell et al., 1994) for the HGA. The MM5 simulation results currently used in the HGA SIP modeling utilized the GOES-satellite skin temperature assimilation technique developed by Dr McNider of University of Alabama Huntsville (personal communication, 2004).

The vertical structure used in the MM5 simulation extended from the surface to the 5000 Pa with 43 sigma levels. The first layer had a vertical thickness of approximately 34 m. The initial and boundary conditions for the MM5 simulation were obtained from three-hourly EDAS analyses, available from NCAR. Sea surface temperatures were extracted from the EDAS analyses. Details are available through TCEQ and Texas A&M and the description of model physics options used can be found in Nielsen-Gammon (2001, 2002). We have used these data as the base meteorological inputs for anthropogenic emissions processing and air quality modeling to provide wind, moisture, temperature fields, parameters determining atmospheric mixing, and deposition velocities.

2.3. Anthropogenic emission inventory

Anthropogenic emissions were derived from the Texas Emissions Inventory (EI) provided by TCEQ. TCEQ's Texas EI was processed through particular steps to provide

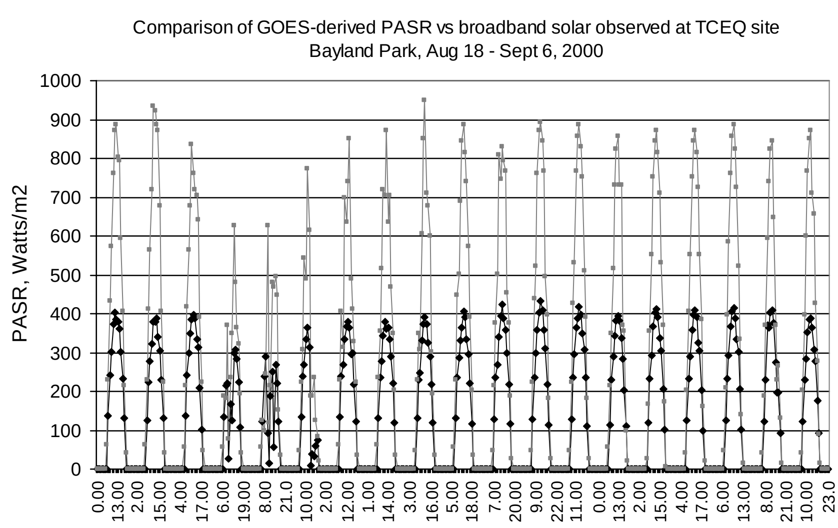

Fig. 6. A sample time series comparison between the GOES-derived PSAR values (darker line) and the broadband solar measurements (grey line) at the Baryland Park TCEQ solar radiation site.

Table 1
Texas emissions inventory provided by TCEQ for the study of TexAQS 2000

Source	EI sub-category	Emission type
Area/nonroad mobile	Texas area	Peak ozone day
	Texas nonroad	Peak ozone day
	Louisiana emissions	Peak ozone day
	Off-shore (shipping lanes)	Peak ozone day
	Elevated ship emissions	Peak ozone day
Point	Texas EGU (electricity generation unit)	Hourly and peak ozone day
	Texas NEGU (Non-EGU)	Peak ozone day
	Louisiana EGU	Hourly and annual average day
	Louisiana NEGU	Annual average day
	Off-shore (platforms)	Peak ozone day
	Texas upset and additional	Hourly and peak ozone day
On-road mobile	MOBILE5 and 6 outputs	Hourly and link-based

model-ready emissions inputs for air quality modeling (Kim and Byun, 2003). TCEQ, Environ, the University of Texas, and others have implemented specific emissions processing methods for building the Texas EI (ftp://www.tnrcc.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Modeling/EI/) used for the HGA state implementation plan modeling studies. In particular, the inventory data includes the Houston-Galveston Ship Channel point-source speciated VOC emissions and road network link-based hourly mobile emissions for the HGA area estimated with MOBILE6 revised by TTI (The Texas Transportation Institute). Table 1 provides Texas Emissions Inventory categories for area and point source emissions inventories provided by TCEQ.

3. Calculation methods

3.1. Biogenic emissions processing with LANDSAT based land cover data

GloBEIS3 (Guenther et al., 2000; Yarwood et al., 1999, 2002) was used here to process biogenic emissions for the Houston-Galveston air quality modeling. GloBEIS3 requires land use land cover and vegetation data for the estimation of biogenic emissions. One of the main purposes of this study is to compare GloBEIS3 results obtained with the TCEQ-LULC and TFS-LULC data. To isolate the impact of the new LULC data on the biogenic emissions estimates, we utilized the observed temperature and GOES-derived radiation fields that are not dependent on the LULC data as described above.

The GloBEIS3 is composed of the Microsoft ACCESS codes. It provides an easy-to-update environment for input data such as land use land cover and meteorological data (Yarwood et al., 2002). As a first step, to calculate the fractional LULC coverage by cell in a model domain, we digitized the TFS-LULC data which is in GIS shape file format using the MIMS spatial allocator (Eyth and Hanisak, 2003). Table 2 presents the TFS-LULC classes used in this study and the new LC codes assigned for the emissions processing. NO emission rates and emission factors which depend on the LULC class are also presented in the table.

To estimate emissions of the isoprene and other lumped VOC species, such as total monoterpenes and higher carbon volatile organic compounds, the detailed vegetation species for the LULC class and the leaf mass density (LMD) data by species must be specified in the internal look-up table. Currently, TCEQ and Environ do not provide such data for monoterpenes. Compared to the TCEQ-LULC data, the TFS-LULC data uses quite different number of LULC classes and thus it is difficult to directly use the vegetation data for the TCEQ-LULC data for the biogenic emissions processing.

On the other hand, the US Forest Service recently conducted a vegetation survey to develop the LMD data for the HGA using a similar land use land cover classes as in the TFS-LULC data. Therefore, we decided to combine the TFS-LULC data with the LMD data from US Forest Service (USFS) instead of the vegetation and LMD data from Wiedenmyer et al. (2001), which was used with the TCEQ-LULC data. The detailed vegetation and LMD data for each TFS-LULC class are tabulated in the Appendix.

Fig. 7 describes the processing steps used to estimate biogenic emissions with the GloBEIS3 with the TFS-LULC and USFS LMD data for HGA. Using the 'data entry'

Table 2
The mapping table for TFS-LULC classes and the leaf mass density data used for GloBEIS3 processing

No.	TFS-LULC class	LC code	NO emission ($\mu g\ m^{-2}\ h^{-1}$)	US forest service LMD class
1	Barren/urban impervious	13590	4.5	Urban built
2	Forest broadleaf	37590	4.5	Northern broadleaf
3	Forest coniferous	37591	4.5	Northern coniferous
4	Forest mixed	37592	4.5	Northern mixed
5	Grass	22590 22591	100.0 100.0	Northern agriculture/range—Chambers, Harris, Liberty, Montgomery and Waller. Southern agriculture/range—Brazoria, Ft. Bend and Galveston
6	Impervious	13590	4.5	Urban built
7	Water	40000	0.0	Water

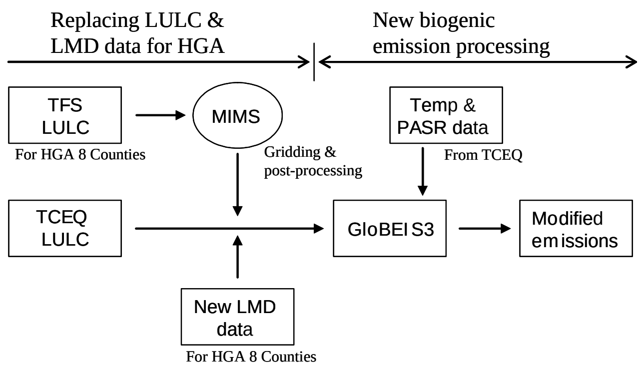

Fig. 7. Modified estimation of GloBIES3 biogenic emissions utilizing new land use (LU), land cover (LC), and leaf mass density (LMD) data for the Houston-Galveston area (HGA). MIMS stands for a GIS-shape file processor supported by US EPA.

module in GloBEIS3, the new LC codes and detailed leaf biomass data were incorporated to build a new input database for the biogenic emission processing.

3.2. CAMx modeling

3.2.1. CAMx-ready emissions data processing

TCEQ has prepared the Texas emissions inventory (EI) used for the Houston-Galveston Area (HGA) state implementation plan (SIP) modeling studies. The inventory data are processed through the EPS2 (Emissions Preprocessing System Version 2) system (Environ, 2002b), GloBEIS3, and EPA's MOBILE6 modified by Texas Transportation Institute (TTI). The emissions data are used in the CAMx air quality model (Environ, 2002a) to assess the efficacy of the emissions control strategies for the Houston-Galveston area.

3.2.2. CAMx-simulations

We performed air quality simulations using CAMx with the identical model options used for the air quality modeling for the development of SIP. It utilizes the Lambert conformal map projection, the piece-wise parabolic method (PPM) numerical advection and vertical eddy diffusion algorithms for transport, and the Carbon Bond version 4 (CB-4) as the chemistry mechanism. The simulations were performed using the same anthropogenic emissions but with the different biogenic emissions estimated with the two different land cover datasets as described above. The results were compared with the TCEQ's Continuous Air Monitoring Site (CAMS) monitoring data.

4. Results and discussion

4.1. Effects of land use and land cover data on biogenic emission estimates

New biogenic emissions data processed with the TFS-LULC and LMD data for the HGA were compared with those based on the TCEQ-LULC data. Outside the HGA, the existing biogenic emission data were used without modification. Fig. 8 compares domain-wide biogenic emissions estimated with GloBEIS3 based on the two land cover data for the HGA, which covers less than the half of the total land area for the modeling domain. While the NO emissions with the GloBEIS3 estimations with TFS-LULC presented ~20% increased biogenic NO emissions, CO and ISOP (isoprene) emissions were decreased by ~10%. Since NO emissions depend on land cover types rather than the vegetation itself in GloBEIS3, the land coverage and the associated NO emissions factor by LC class for TFS-LULC presents different NO emissions for the area. For CO and ISOP emissions, not only land cover changes in TFS-LULC but also the new LMD applied affected on the GloBEIS3 estimations. The differences become especially larger during afternoons when PASR and temperature are increased. No biogenic ISOP emissions are present before sunrise and after sunset when PASR is not available. Emissions rates for all species during the last three days (August 30–September 1, 2000) are higher than other days because the higher maximum temperatures were observed during the last period of the episode.

For the study, we used GloBEIS version 3.1 for the CB-4 mechanism as it is processed for the HGA SIP, in which no speciation factor for terpenes from the lumped VOC is available, and therefore we have focused on the speciation of precursor emissions for gas-phase chemistry reactions. Although CO is not active species, it was used to explain the difference between two LULC data.

Fig. 9 depicts snapshots of NO emissions for August 25th, 2000 at 12 CST and compares the spatial changes in the emission rates for the domain. It should be noted that emission rates were the same outside the Houston-Galveston 8 county areas because no modification was made on the land cover and LMD data. The GloBEIS3 estimation with TFS-LULC shows decreased NO emission rates in Fort-Bend County. We suspect that the large difference in the NO emissions between Fort-Bend County and its neighboring counties in the east and south is due to the artifact in the original TCEQ-LULC database, which partially relied on county-based GIS surrogates. With the TFS-LULC data, the NO emissions are more or less evenly distributed for the 8-county areas. Higher NO emissions in Waller County, the southern part of the Harris County, Chambers County and Galveston County are caused by the high fraction of grass land cover. Except for the Fort-Bend County, the biogenic NO emissions with TFS-LULC are higher than those estimated with the original land cover data. With the more detailed land cover data, the NO emissions from the other seven counties are considerably increased.

In general, the biogenic CO and ISOP emissions (Figs. 10 and 11) have decreased with the TFS-LULC data. In particular, the areas around Montgomery and Liberty Counties show significantly lower CO emissions

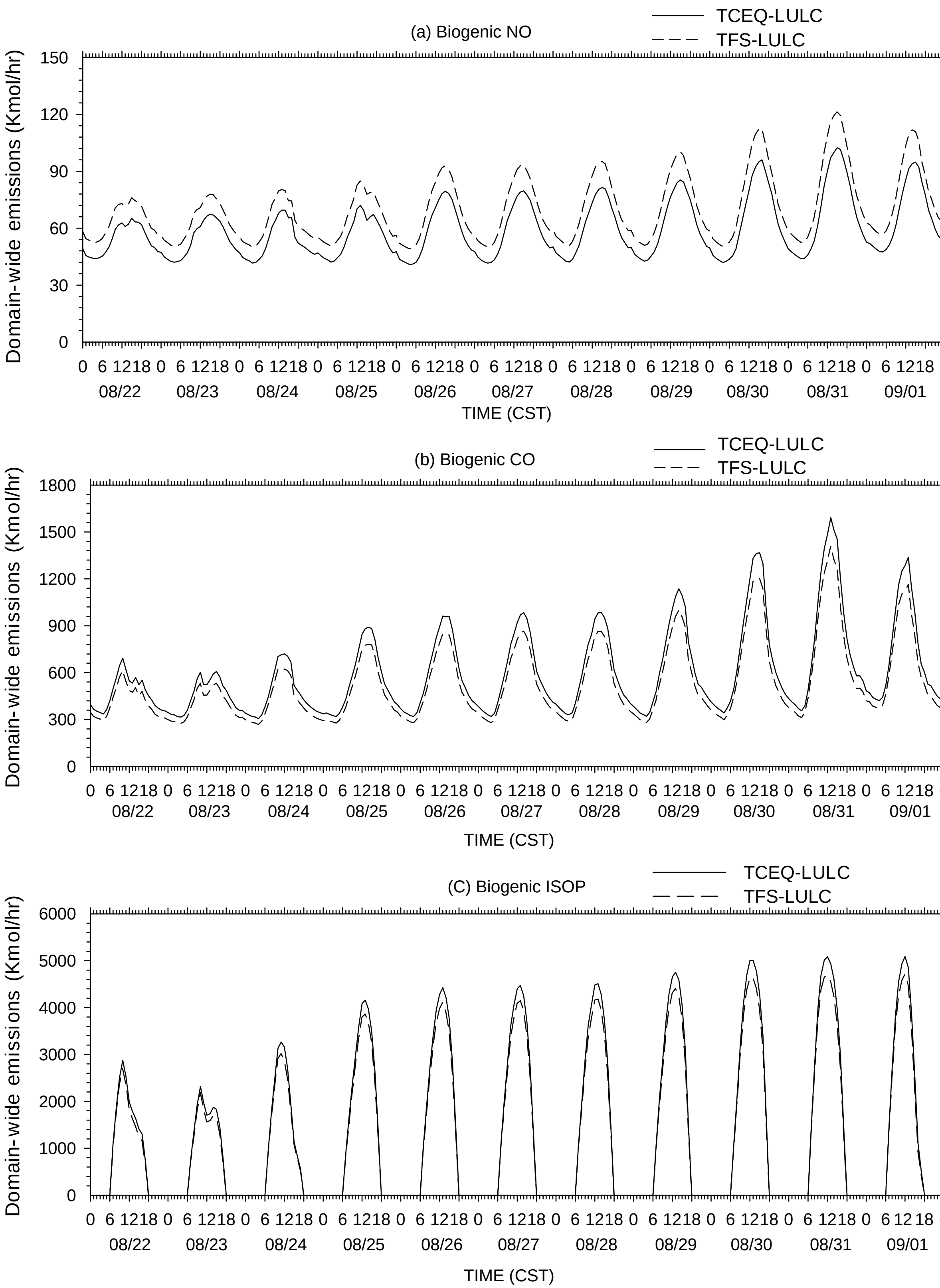

Fig. 8. Diurnal cycles of biogenic emissions estimated with the TCEQ-LULC and TFS-LULC data for (a) NO, (b) CO, and (c) ISOP.

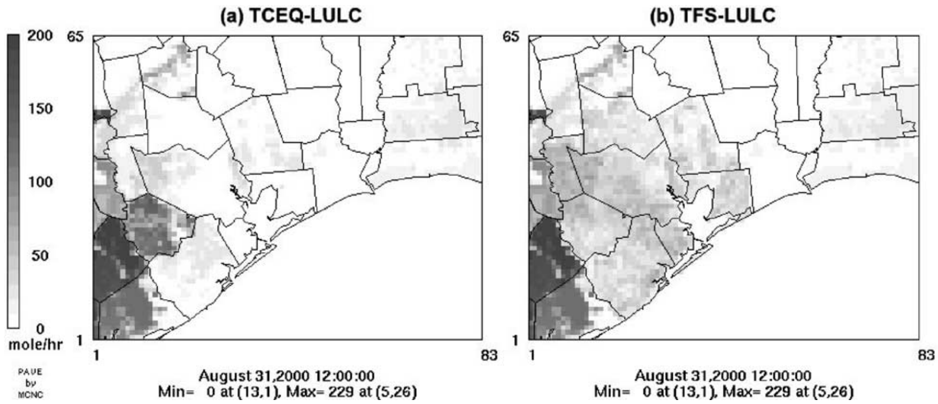

Fig. 9. Spatial comparison of NO emission rates estimated using TCEQ-LULC and TFS-LULC data for August 31, 2000 at 12 CST.

with new land cover data than the TCEQ-LULC case. The shape of the area of the biogenic emissions decrement looks similar to the Mixed Forest class in TFS-LULC. Other areas of VOC emissions decrease are in northern HGA (Montgomery County) along with the forest area (not shown). Also, there is a general decrease in the biogenic VOC emissions in the areas where suburban/urban development has occurred.

The areas with the most decreased ISOP emissions are located inside Liberty County, followed by Montgomery County. This is where the 'Bottom Oaks' tree species are located in the TCEQ-LULC database. It is suspected that the difference at this area may be due to the insufficient tree survey details available to match with the satellite-driven TFS-LULC classification. Most of the areas show decreased emission rates but we can find a few spots at which ISOP emissions increased with the TFS-LULC data, especially in Chambers County. No large changes in Harris County and around the Ship Channel area are observed.

Isoprene emissions are mostly from biogenic sources. Differences shown here affect ozone concentrations over the downwind areas of the high isoprene emissions areas. On the other hand, the contributions from the biogenic emissions of NO and CO are relative small (~5 and 8% of the domain totals) compared to their anthropogenic counterparts such as mobile, area, and point sources. Therefore, it is expected that the changes in the biogenic NO and CO emissions have negligible impact on air quality simulation results. Fig. 12 compares the biogenic emissions of NO, CO and ISOP relative to the total emissions over the model simulation domain.

4.2. Sensitivity of ozone predictions to the different biogenic emissions

To estimate the impact of biogenic emission changes corresponding to the land cover changes on ambient ozone concentrations in air quality model predictions, we have prepared two emission inputs to CAMx with GloBEIS3 using TCEQ-LULC and TFS-LULC data coupled with TCEQ's anthropogenic emissions inventory processed through EPS2. CAMx simulation results were compared with TCEQ Chemistry Air Monitoring Site (CAMS) observations (Fig. 13a–c).

Since there were appreciable changes in biogenic emissions with favorable weather conditions (i.e. high temperature and PASR) for the last three days, we present snapshots of the differences in O_3 and ISOP concentrations at 15 CST of each day and time series analysis at the two

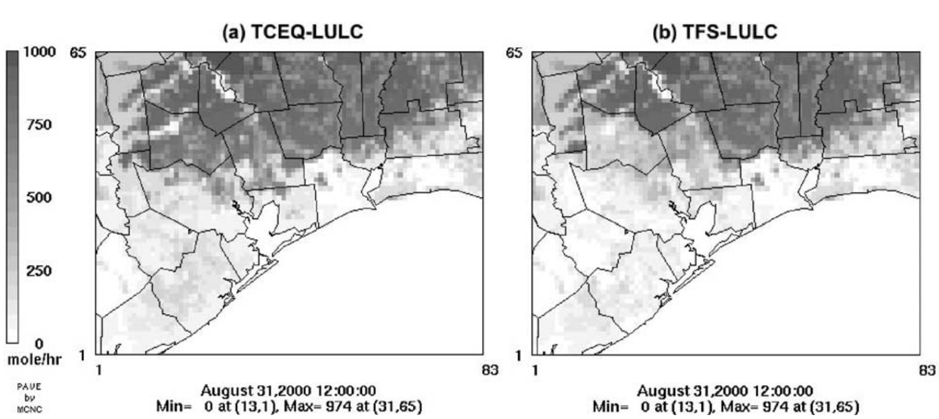

Fig. 10. Spatial comparison of CO emission rates estimated using TCEQ-LULC and TFS-LULC data for August 31, 2000 at 12 CST.

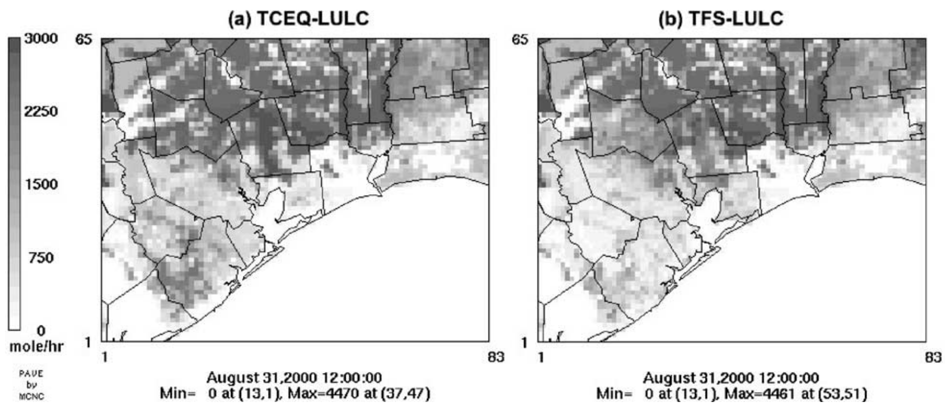

Fig. 11. Spatial comparison of isoprene (ISOP) emission rates estimated using TCEQ-LULC and TFS-LULC data for August 31, 2000 at 12 CST.

sites on which both O_3 and ISOP measurements were made during the span.

4.2.1. Differences in ozone concentrations

The region affected by the changes in the biogenic emissions is determined by the wind direction. For example, on August 25th and 29th, the areas of large differences occur away from the Ship Channel industry and therefore the regional peak ozone values are not changed significantly. However, on August 30th, 31st, and September 1, areas of perceivable ozone reduction are shown near the Ship Channel as well (Fig. 14).

CAMx simulations showed more than 10 ppb (up to 25 ppb) differences between the two biogenic emissions. As ISOP emissions decreased with TFS-LULC, O_3 concentrations also decreased for these days. The areas of large difference are located in the southern part of Harris County, Fort-Bend County and Galveston County during afternoons.

Predicted O_3 concentrations compared to measurements at La Porte and Clinton (not shown) indicated that O_3 concentrations decreased by ~ 10 ppb for the last three days and maximum differences occurred around 15 CST (3 PM) as both simulation and observation reached the daily peaks. Also, there are large areas of lower ozone concentrations predicted with the new biogenic emissions in Port Bend County and Brazoria County. The southwest area is NO_x rich, because there is a large amount of NO_x from mobile sources and from the Parish power plant in that area. Therefore, the ozone originating in that area is probably formed in VOC-limited conditions, so changes in the VOCs (isoprenes) from biogenics will have a substantial effect. Occasionally, August 29 and 31 in particular, it seems to show that the footprint of the Parish power plant plume is more pronounced in the run with biogenic emissions with TFS-LULC.

As for the ozone changes in the Baytown area, we think that those are due to the problem with the TFS-LULC's characterization of the Trinity River bottomland forests in Liberty and Chambers Counties. In Chambers County, the satellite-driven data seems to account for the unique vegetation distribution in the bottomland forest, but in Liberty County it does not. The TCEQ data set categorized the entire bottomland forest as a unique LULC class with a unique vegetation distribution, whereas the new dataset does not. The result is that the new data shows much lower isoprene emissions in Liberty County than the TCEQ data, but similar or greater emissions in Chambers County. Note that the NOAA and DOE aircraft measurements of VOCs taken during TexAQS seem to indicate that the isoprene emissions really are higher along the bottomlands than in the adjacent areas, but not in the Houston Ship Channel area, so this effect seems to be verified by the VOC data (Daum et al., 2003; Ryerson et al., 2003).

4.2.2. Differences in ambient isoprene concentration

By comparing simulated ISOP concentrations, we can evaluate the effects of land cover changes on each county. In Fig. 15, ISOP concentrations in Harris and Galveston counties remain relatively the same, while those in other encompassing counties decreased except in Chambers County. The diurnal cycle plots of ambient isoprene concentrations at the two measurement sites usually show two peaks a day by morning and by afternoon (see Fig. 16). GloBEIS3 estimates ISOP emissions to start from 7 AM (CST) when the PASR becomes available. At this time, photochemistry is not yet active. Therefore, we see as much difference in the ambient isoprene concentrations between

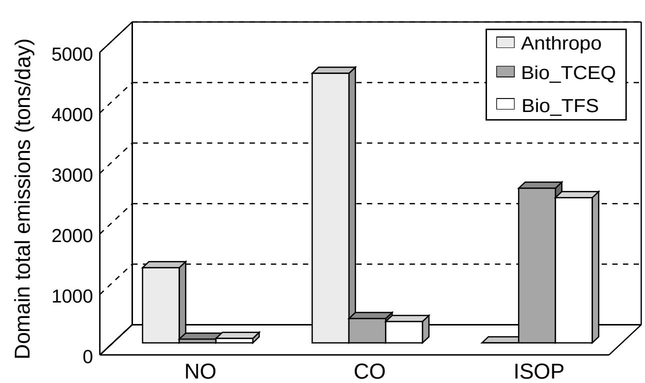

Fig. 12. Comparison of domain total NO, CO and ISOP emissions from anthropogenic and biogenic sources during the period.

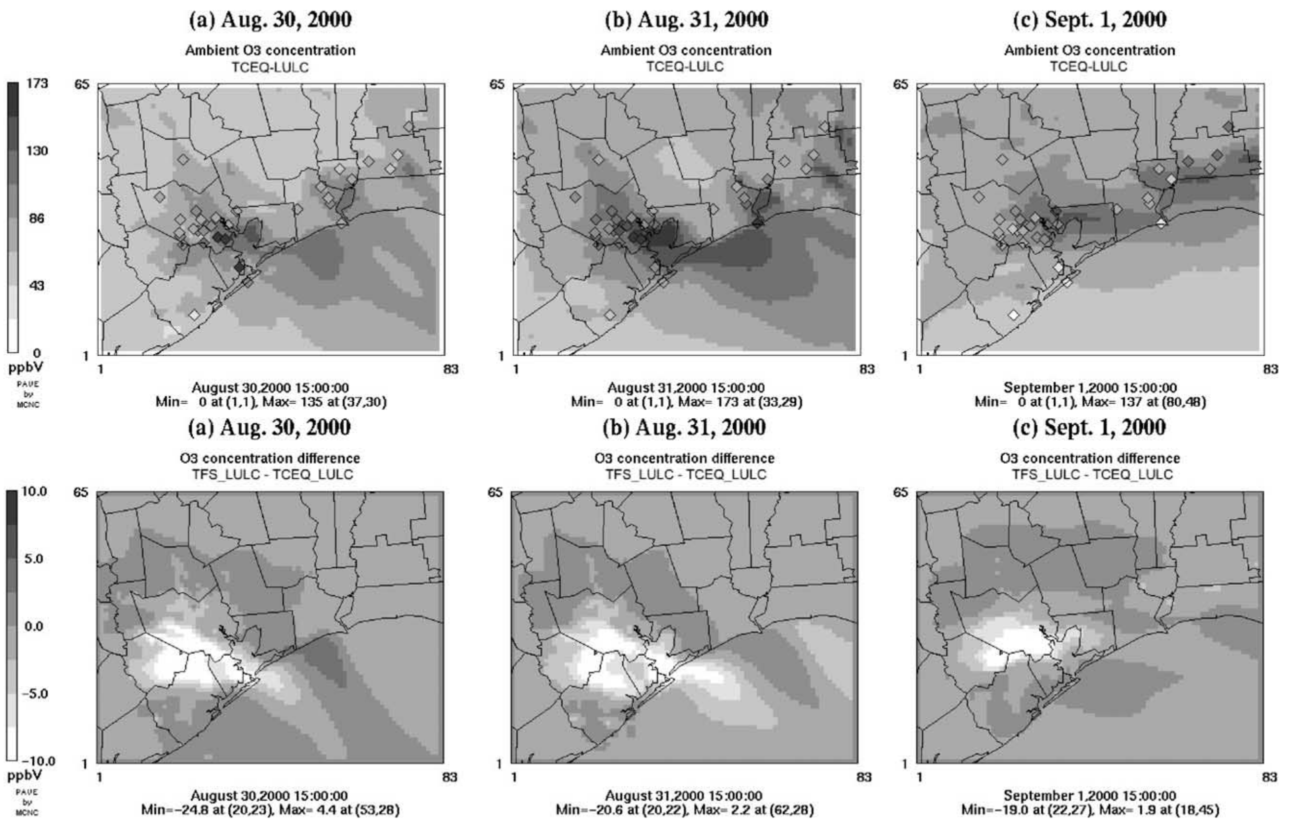

Fig. 13. Ozone concentration distributions predicted by CAMx with the TCEQ-LULC and the difference with those with the TFS-LULC for (a) and (d) August 30, (b) and (e) August 31, and (c) and (f) September 1, 2000 at 15 CST. Diamond symbols represent observed surface ozone concentrations at the CAMS sites.

the two land cover data as in the biogenic emissions. However, the difference begins to decrease as photochemistry becomes active and isoprene molecules are used for the ozone formation.

Another difference in the peak appears around 3 PM when ISOP emissions reach the daily maximum. It is suspected that most of the fresh NO_x molecules in the atmosphere have been consumed to form terminal products or sequestered as PAN. Therefore, the propagation of the OH radical diminishes considerably and further ozone formation is essentially stopped at that moment. We also suspect that the vertical eddy diffusivity in the CAMx model may be too low to represent the rigorous convective boundary layer mixing. This can be seen from the time series plots for the La Porte and Clinton sites (Fig. 16). We will investigate the causes of these peculiar temporal characteristics in future sensitivity studies.

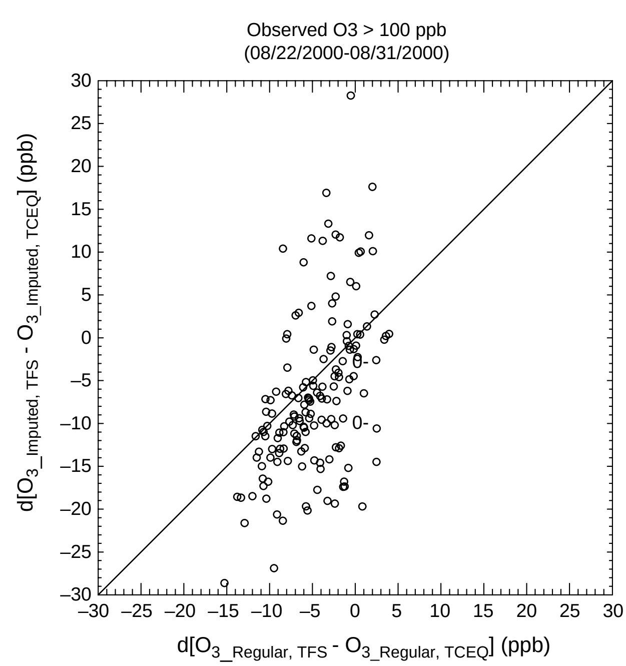

Fig. 14. Sensitivity test on ozone concentrations changes for the CAMx simulations with TFS-LULC and TCEQ-LULC data. The differences in simulated ozone concentrations for regular and imputed emissions inventories were compared only when the observed values exceeded 100 ppb at the CAMS sites during the simulation period.

5. Conclusions

We have analyzed the impact of using satellite-derived land use and land cover data on biogenic emissions estimates. The field plot information, combined with satellite estimates of tree cover (stratified by the land use type specified in the new dataset) and detailed leaf mass density data for the Houston-Galveston area developed by

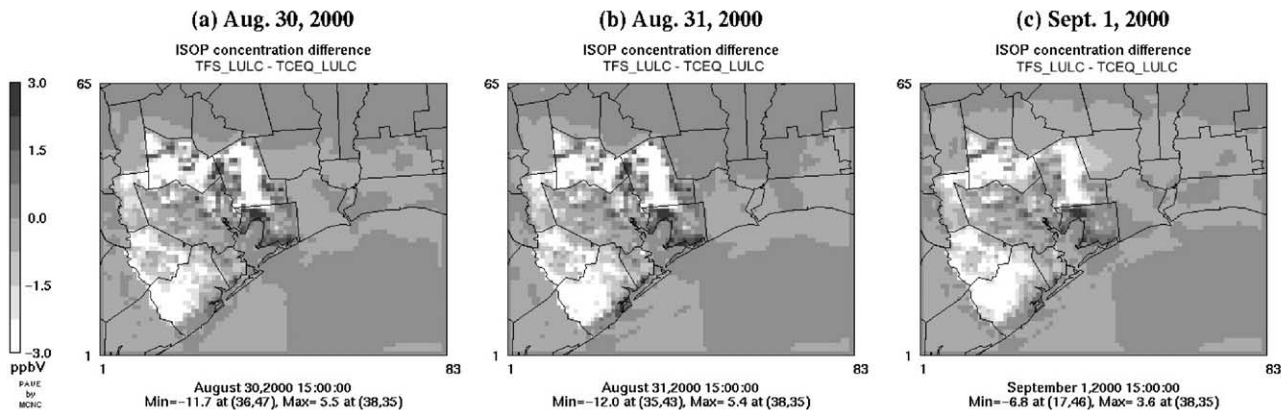

Fig. 15. ISOP concentration differences in CAMx simulations with different biogenic emissions using the two LC data for (a) August 30, (b) August 31, and (c) September 1, 2000 at 15 CST.

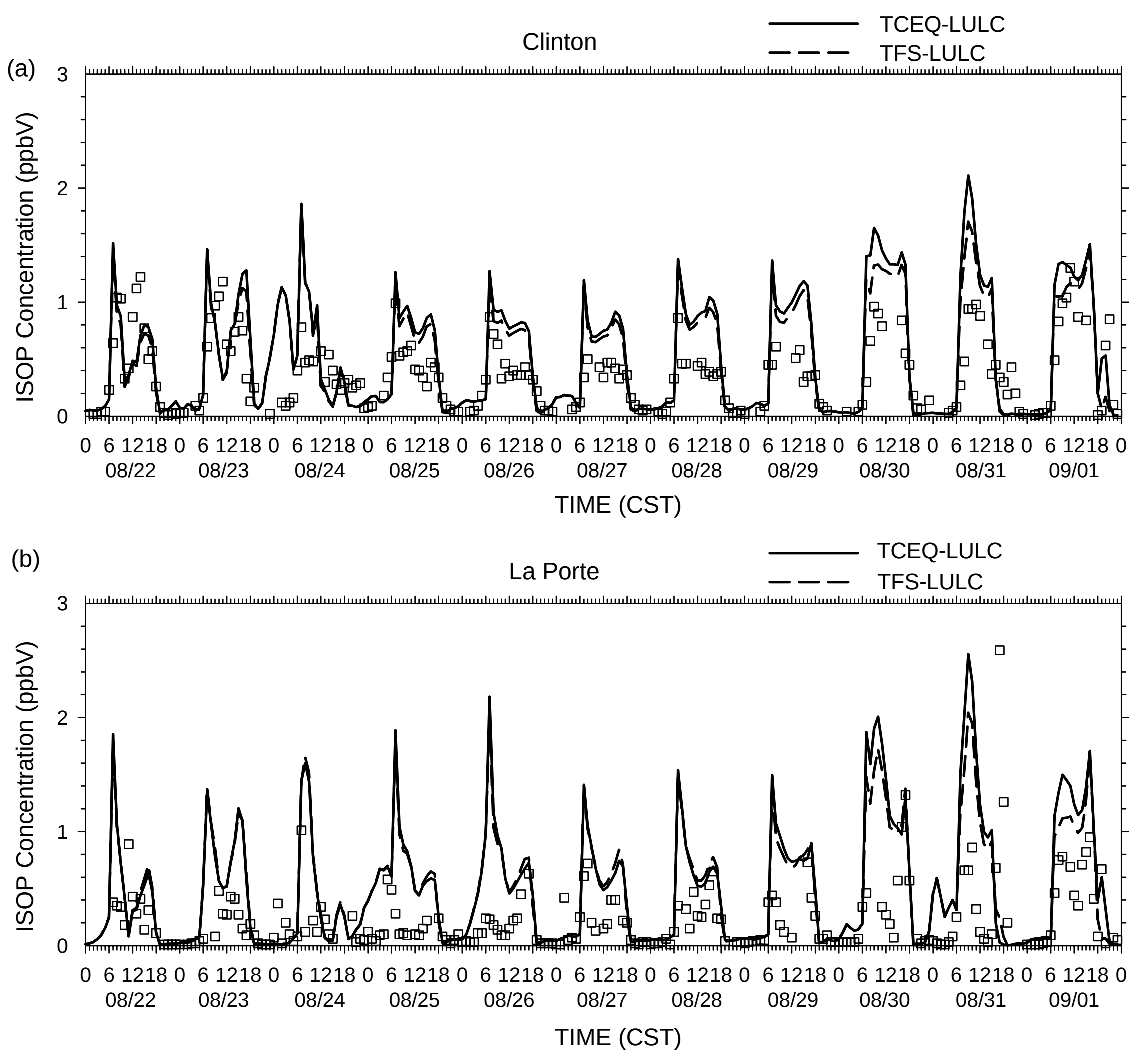

Fig. 16. Predicted ISOP concentrations using two LC data compared to those observed at (a) Clinton and (b) La Porte sites during the period of August 22–September 1, 2000.

the US Forest Service were used to estimate changes in the biogenic emissions. Air quality simulations were performed using both the original and the new biogenic emissions inputs.

The results showed that the TFS LANDSAT data resulted in higher amounts of biogenic NO emissions and lower isoprene and CO emissions than those from the TCEQ-LULC data. Because the contributions of biogenic NO and CO emissions to the anthropogenic counterparts were relatively small, the differences in the isoprene emissions amount affected air quality simulation results most. There were considerable differences in biogenic isoprene emissions, subsequently, the ozone predictions were affected up to 10 ppb, but the magnitudes varied each day depending on the upwind or downwind positions of the biogenic emission sources relative to the anthropogenic NO_x and VOC sources. Although the assessment had limitations such as the heterogeneity of the spatial resolutions between the two datasets, the study highlighted the importance of biogenic emissions uncertainty on air quality predictions. However, the study did not allow extrapolation of the directional changes in air quality corresponding to the changes in the LULC because the two datasets use vastly different LULC category definitions and uncertainties in the vegetation distributions.

This research focused only on the direct effects of using high resolution land cover data on the biogenic emission estimates and subsequent ozone predicted by CAMx. Land use and land cover significantly affect land-surface energy budget and meteorological conditions. Therefore, in order to account for the full effects of LULC changes on the local air quality, one must first address changes in the meteorological conditions. We have started evaluating methods utilizing the LANDSAT derived LULC data in the MM5 meteorological simulations. In addition, a 1990 land cover data set is under development using past LANDSAT data to evaluate the changes during the last decade. In the near future, we will assess the overall effects of different land use inputs on meteorological conditions, biogenic and anthropogenic emissions, and air quality.

Acknowledgements

The research described in this article has been funded in part by the Houston Advanced Research Center (HARC) H17: Modeling Effects of Land Use/Land Cover Modifications on the Urban Heat Island Phenomena in Houston, Texas, and the United States Environmental Protection Agency through Grants R-82906801 and R-83037701 to the University of Houston. However, it has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. This work was also funded in part by the Texas Air Research Center and their contributions are acknowledged.

Appendix. Leaf biomass for vegetation species by LULC class in the Houston-Galveston area from US forest service

Genus	Leaf biomass (g/m^2)
LC code: 13590, urban built	
Quercus	7.86
Sapium	2.49
Salix	1.80
Juniperus	1.48
Thuja	1.07
Ulmus	0.51
Carya	0.45
Liquidambar	0.30
Celtis	0.21
Pinus	0.05
Total	16.22
LC code: 22590, Northern agriculture/range	
Quercus	12.85
Sapium	2.97
Liquidambar	2.52
Ulmus	1.47
Nyssa	1.05
Zanthoxylum	0.63
Magnolia	0.53
Other	0.00
Total	22.02
LC code: 22591, Southern agriculture/range	
Sapium	8.54
Ulmus	4.47
Quercus	3.03
Carya	1.86
Celtis	0.99
Bumelia	0.80
Pinus	0.27
Acer	0.20
Crataegus	0.10
Fraxinus	0.09
Other	0.00
Total	20.35
LC code: 37590, Northern broadleaf	
Quercus	63.45
Pinus	52.24
Sapium	25.02
Taxodium	23.58
Liquidambar	13.00
Ulmus	11.95
Fraxinus	8.76
Nyssa	3.53
Acer	3.20
Juniperus	2.53
Celtis	1.81
Juglans	1.12
Platanus	1.00
Carya	0.93
Magnolia	0.89
Carpinus	0.85
Diospyros	0.64
Crataegus	0.53
Ilex	0.52
Planera	0.44
Persea	0.32

(*continued on next page*)

Genus	Leaf biomass (g/m^2)
Cornus	0.14
Salix	0.10
Other	0.00
Total	216.55
LC code: 37591, Northern coniferous	
Pinus	300.02
Quercus	120.91
Sapium	16.38
Fraxinus	13.83
Carya	10.37
Carpinus	8.46
Ilex	8.37
Ulmus	7.68
Acer	4.70
Liquidambar	4.66
Persea	2.33
Celtis	2.21
Nyssa	1.38
Ostrya	0.01
Total	501.31
LC code: 37592, Northern mixed	
Quercus	96.86
Pinus	81.34
Carpinus	21.25
Liquidambar	19.19
Carya	15.42
Ulmus	12.31
Salix	11.60
Acer	10.38
Sapium	5.98
Fraxinus	5.71
Ostrya	3.23
Taxodium	2.86
Nyssa	2.42
Morus	0.47
Prunus	0.40
Other	0.00
Total	289.42

References

Byun, D.W., Kim, S.-T., Cheng, F.-Y., Kim, S.-B., Cuclis, A., Moon, N.-K., 2003. Information infrastructure for air quality modeling and analysis: application to the Houston-Galveston ozone nonattainment area. J. Environ. Inform. 2 (2), 38–57.

Daum, P.H., Kleinman, L.L., Springston, S.R., Nunnermacker, L.J., Lee, Y.-N., Weistein-Lloyd, J., Zheng, J., Berkowits, B., 2003. A comparative study od O3 formation in the Houston urban and industrial plumes during the 2000 Texas Air Quality Study. J. Geophys. Res. 108, 4715.

Environ, 2002a. User's Guide to the Comprehensive Air Quality Modeling System with Extensions (CAMX), Version 3.10.

Environ, 2002b. User's Guide to the Global Biosphere Emissions and Interactions System (GloBEIS3), Version 3.0.

Eyth, A., Hanisak, K., 2003. The MIMS spatial allocator: a tool for generating emissions surrogates without a geographic information system. Proceedings, 12th International Emission Inventory Conference, San Diego, April 29–May 1, 2003 (available at http://www.epa.gov/ttn/chief/conference/ei12/).

Funk, T.H., Stiefer, P.S., Chinkin, L.R., 2001. Development of Gridded Spatial Allocation Factors for The State of Texas, Final Report 3, STI-900570-2114-FR3, Prepared for Texas Natural Resource Conservation Commission, August 31, 2001 (www.tnrcc.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Contract_Reports/EI/).

GEM, 2003. Satellite Data Processing for Land Use and Cover Type Classification, Final Report, TFS-3-027, Prepared for the Texas Forest Service, July 8, 2003.

Geron, C.D., Guenther, A.B., Pierce, T.E., 1994. An improved model for estimating emissions of volatile organic compounds from forests in the eastern United States. J. Geophys. Res. 99 (D6), 12773–12791.

Grell, G.A., Dudhia, J., Stauffer, D.R., 1994. A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Technical Note, NCAR/TN-398 + STR, 122 pp.

Guenther, A.B., Zimmerman, P.R., Harley, P.C., Monson, R.K., Fall, R., 1993. Isoprene and monoterpene emission rate variability: model evaluations and sensitivity analyses. J. Geophys. Res. 98D, 12609–12617.

Guenther, A., Zimmerman, P., Wildermuth, M., 1994. Natural volatile organic compound emission rate estimates for US woodland landscapes. Atmos. Environ. 28, 1197–1210.

Guenther, A., Geron, C., Pierce, T., Lamb, B., Harley, P., Fall, R., 2000. Natural emissions of non-methane volatile organic compounds, carbon monoxide, and oxides of nitrogen from North America. Atmos. Environ. 34, 2205–2230.

Kim, S., Byun, D.W., 2003. Prototyping the Texas emissions inventory preparation system for the SMOKE system. 12th International Emission Inventory Conference, San Diego, April 29–May 1, 2003 .http//www.epa.gov/ttn/chief/conference/ei12/poster/kim.pdf.

Kinnee, E., Geron, C., Pierce, T., 1997. United States land use inventory for estimating ozone precursor emissions. Ecol. Appl. 7, 46–58.

Nielsen-Gammon, J.W., 2001. Initial modeling of August 2000 Houston-Galveston ozone episode. Report to the Technical Analysis Division, Texas Natural Resource Conservation Commission. December 2001 (available at http://www.tnrcc.state.tx.us/air/aqp/airquality_contracts.html#met01).

Nielsen-Gammon, J.W., 2002. Meteorological Modeling for the August 2000 Houston-Galveston Ozone Episode: METSTAT Statistical Evaluation and Model Runs from March–June 2002. Report to the Technical Analysis Division, Texas Natural Resource Conservation Commission. June 2002 (available at Department of Atmospheric Sciences, 3150 TAMU, College Station, TX 77843-3150).

Nielsen-Gammon, J.W., Walter, K., Epifanio, C., Chadwick, V., 2004. The surprising dynamics of the Houston urban sea breeze, Interactive Symposium on the Nexus of the Coastal and Urban Environments, American Meteorological Society, Seattle, WA.

Nowak, D.J., Civerolo, K.L., Rao, S.T., Sistla, G., Luley, C.J., Crane, D.E., 2000. A modeling study of the impact of urban trees on ozone. Atmos. Environ. 34, 1601–1613.

Oke, T.R., 1987. Boundary Layer Climates, second ed. Methuen, London.

Pierce, T., Geron, C., Bender, L., Dennis, R., Tennyson, G., Guenther, A., 1998. The influence of increased isoprene emissions on regional ozone modeling. J. Geophys. Res. Atmos. 103, 25611–25629.

Pinker, R.T., et al., 2003. Surface radiation budgets in support of the GEWEX continental-scale international project (GCIP) and the GEWEX Americas prediction project (GAPP), including the North American land data assimilation system (NLDAS) project. J. Geophys. Res. 108 (D2), 8844.

Ryerson, T.B., et al., 2003. Effect of petrochemical industrial emissions of reactive alkenes and NOx on tropospheric ozone formation in Houston, Texas. J. Geophys. Res. 108 (D8), 4249.

Taha, H., Chang, S., Akbari, H., 2000. Meteorological And Air Quality Impacts Of Heat Island Mitigation Measures In Three US Cities. LBNL-44222. Lawrence Berkeley National Laboratory, Berkley, CA.

Texas Commission on Environmental Quality, 2002. Attachment 3—Emissions Inventory Development and Modeling for the August 25–September 1, 2000 Episode November 15, 2002 (ftp.tnrcc.state.tx.us/pub/OEPAA/TAD/Modeling/HGAQSE/Modeling/Doc/TSD_PHASE1/attachment3-emissions_inventory.pdf).

Vizuete, W., Junquera, V., McDonald-Buller, E., McGaughey, G., Yarwood, G., Allen, D., 2002. Effects of temperature and land use on predictions of biogenic emissions in Eastern Texas, USA. Atmos. Environ. 36, 3321–3337.

Wiedenmyer, C., Guenther, A., Estes, M., Strange, I.W., Yarwood, G., Allen, D.T., 2001. A land use database and examples of biogenic isoprene emission estimates for the state of Texas, USA. Atmos. Environ. 35, 6465–6477.

Yarwood, G., Wilson, G., Emery, C., Guenther, A., 1999. Development of GloBEIS—A State of the Science Biogenic Emissions Modeling System. Final Report, Prepared for Texas Natural Resource Conservation Commission.

Yarwood, G., Wilson, G., Shepard, S., 2002. User's Guide to the Global Biosphere Emissions and Interactions System (GloBEIS) Version 3; ENVIRON International Corporation; Guenther A., National Center for Atmospheric Research, 2002.

Daewon Byun is presently a professor in the Department of Geosciences, with joint appointment in the Department of Chemistry, University of Houston (UH), and the director for the UH Institute for Multi-dimensional Air Quality Studies (IMAQS). Prior to joining UH, he was the technical leader of the US Environmental Protection Agency (EPA) Models-3 Community Multiscale Air Quality (CMAQ) model development project.

Soontae Kim is a postdoctoral research associate at the IMAQS in UH.